Quality-Control of Upper-Air Soundings for OLYMPEX

PI: Prof. Steven A. Rutledge

Report prepared by Paul Ciesielski

Department of Atmospheric Science
Colorado State University

6 July 2016

1. Introduction

During the 29 October 2015 to 16 January 2016 period, upper-air soundings were collected as part of the OLYMPEX field campaign from the following upper-air sites: the NPOL site, the Environmental Canadian (ENVC) site, two US NWS stations (Quillayute, WA / KUIL and Salem, OR / KSLE), and one Canadian WS station (Port Hardy, BC / CYZT). The location of these five sites is shown in Fig. 1. In addition, several dropsondes were obtained from the NASA-DC8 platform. This report discusses the post-processing of the upsonde portion of the OLYMPEX sounding dataset.

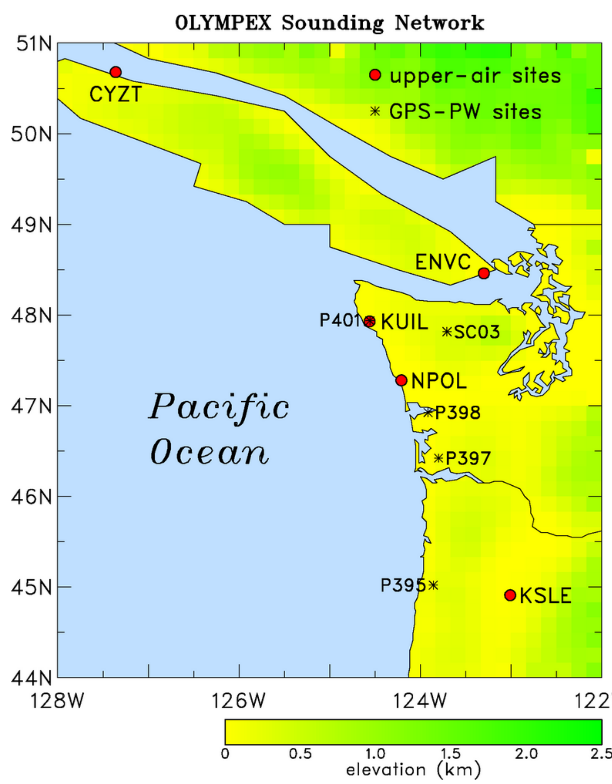

Figure 1. Map showing location of OLYMPEX upper-air sites (red circles) and GPS COSMIC Suominet sites (* symbol) which provided 30-min total-column Precipitable Water (PW) data. Elevation information is indicated by background shading with scale at bottom.

2. Overview of OLYMPEX sounding operations

A visual inventory of the sounding data at the upsonde sites is shown in Fig. 2. Generally speaking the operational sites (i.e., Port Hardy / CYZT, Quillayute / KUIL and Salem / KSLE) had two sonde launches per day (at 00 and 12UTC nominal time), although Quillayute's launch frequency increased to 4/day in January. On the other hand, the ENVC and NPOL sites operated on a variable launch schedule based on weather conditions and scientific mission objectives with up to 8/day soundings during some periods.

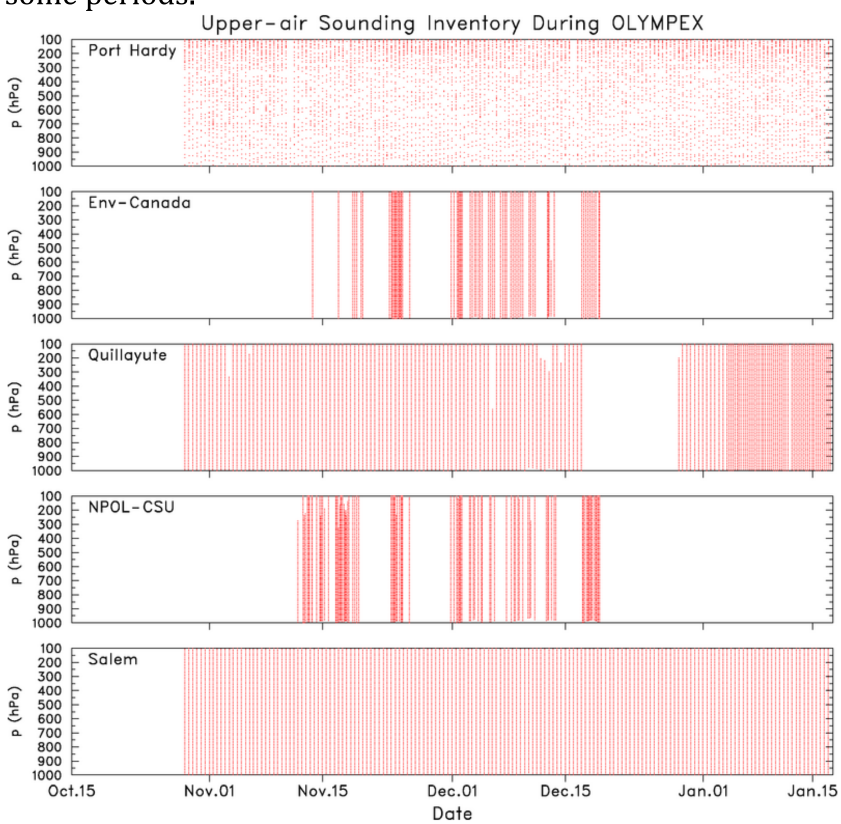

Figure 2. Visual sounding inventory of upper-air data for the period 29 October 2015 to 16 January 2016 based on Level 4 data. Each line of dots represents a successful sonde launch. Higher density of dots with time indicate more frequent launches with up to 8/day observations during certain periods.

High-vertical resolution (hi-res) sonde data (1-2s) was collected from all sites with the exception of Port Hardy, BC (CYZT). The hi-res data from the NWS sites (KUIL and KSLE) were provided by Scot Loehrer of NCAR EOL which routinely archives sounding data from all US NWS operational sites. Unfortunately, no source for hi-res data from CYZT was identified. Instead the GTS-resolution data for this site (as indicated by its diminished vertical resolution in Fig. 2) was obtained from the U. of Wyoming upper-air data archive (http://weather.uwyo.edu/upperair/sounding.html). As noted in Fig. 2, KUIL had no soundings between 12 UTC 17 December to 12 UTC 29 December due to a power outage at this site. Also prior to the 18 November 23UTC sounding at the NPOL site, a significant number of the soundings (23) had missing data (particularly winds) at upper-levels (pressure < 300 hPa)[1]. After this time, use of a different UHF antenna greatly improved the upper-level data reception at the this site. Finally, it should be noted that no baseline ground check was performed on the NPOL soundings prior to 23 November 12UTC due to the lack of adequate desiccant for the GC25 ground check system. Fortunately the absence of this ground check did not appear to adversely impact the quality of the NPOL thermodynamic data.

Table 1 below summarizes the pertinent information for the OLYMPEX upper-air sounding sites. In total 651 upsondes were launched and collected for this experiment. Four of the sites used Vaisala RS92 (VRS92) sondes while the remaining site (KSLE) used Sippican Mark II sondes. The ground station Digicora software used to process the VRS92 sondes in OLYMPEX was version 3.64 which includes corrections to the raw humidity data for a daytime solar radiation dry bias and for a time-lag error due to slow instrument response at cold temperatures. With these corrections the resulting VRS92 humidity profiles in tropical soundings have been found to be extremely accurate (Yu et al. 2015). Wind data from all sites was based on GPS wind retrievals which have an stated accuracy of 0.15 m/s (Vaisala 2014).

The dropsondes were quality-controlled by NCAR EOL and are reported on in a separate document which can be obtained at the following link: (http://data.eol.ucar.edu/codiac/dss/id=503.001).

Site	Station ID	Sonde type	Resolution	No. of soundings	Dates of retrieved data
Port Hardy, BC	71109, CYZT	VRS92	GTS	160	10/29/15 - 01/16/16
Env. Canada	99001, ENVC	VRS92	2 s	71	11/13/15-12/19/15
Quillayute, WA	72797, KUIL	VRS92	1 s	162	10/29/15-01/16/16
NPOL-CSU	99000, NPOL	VRS92	2 s	98	11/11/15-12/19/15
Salem, OR	72694, KSLE	Sip-Mark II	1 s	160	10/29/15-01/16/16
dropsondes	99002, DROP	Vaisala RS904	0.25 s	53	11/12/15-12/19/15

Table 1. Pertinent information for OLYMPEX sounding sites. Resolution refers to the native time resolution of the data. GTS refers to resolution received via GTS (Global Telecommunication System) message.

[1] The operational issues at the NPOL site, noted here, were related to using a Digicora sounding system borrowed from Vaisala which was rushed to the field when several components of the CSU system were stolen en-route to its field deployment.

3. Quality-control procedures

The methodology used to produce a quality-controlled (QC'ed) sounding dataset for OLYMPEX follows that described in Ciesielski et al. 2011. This procedure involves the four stages of processing outlined below.

(1) In this first stage of processing the soundings from the different sites were converted into a single, easily utilized ASCII format (i.e., the GLS format used by NCAR EOL).

(2) Next, the high-vertical resolution (1-2 s) sounding data were passed through a series of automated QC algorithms to systematically detect bad values. For this purpose we used ASPEN (Atmospheric Sounding Processing ENvironment), a software tool developed by NCAR EOL. In addition to removing egregious data based on several objective QC checks (e.g., gross limit, vertical consistency, etc), ASPEN filters the winds, computes geopotential height, smooths pressure and writes out the processed QC'ed sounding data in two convenient formats (i.e., a standard ASCII format used by NCAR EOL and netcdf).

(3) In Level 3 (L3) processing, sonde biases are identified and reduced if possible. While sonde manufacturers are continually striving to improve the accuracy of humidity sensors, water vapor retrieval continues to be the most problematic variable measured (Ciesielski et al. 2012). For identifying humidity biases, we compare the sonde precipitable water (PW) to estimates provided by COSMIC SUOMINET (http://www.suominet.ucar.edu/) GPS PW retrievals. The procedure for computing PW from zenith-path delay (ZPD) data derived from ground-based GPS measurement is described in Wang et al. (2007). To facilitate comparison with sounding data, the 30-min GPS PW data were averaged into hourly bins.

Time series of PW from five GPS sites in the OLYMPEX domain (with locations shown in Fig. 1) exhibit large temporal variability reflecting the changing moisture conditions associated the weather systems which passed over this region (Fig. 3). Differences in their means (shown in parentheses in Fig. 3) are primarily related to differences in the elevation of the GPS instruments, which range from 46 to 2037 m. Since GPS site P401 is collocated with the KUIL NWS sonde launches, we will focus our comparisons on this site.

Figure 4 shows a scatter plot of sonde versus GPS PW for the KUIL site for all times with cotemporaneous observations. As one will note, there is excellent agreement between these two independent measurements of PW with a correlation of 0.98 and a mean bias (sonde – GPS) of +0.07 mm (which is < 0.5% of the mean PW). The handful of cases with PW differences > 2 mm were all associated with very strong winds such that the sonde and GPS instruments were likely sampling different environments.[2]

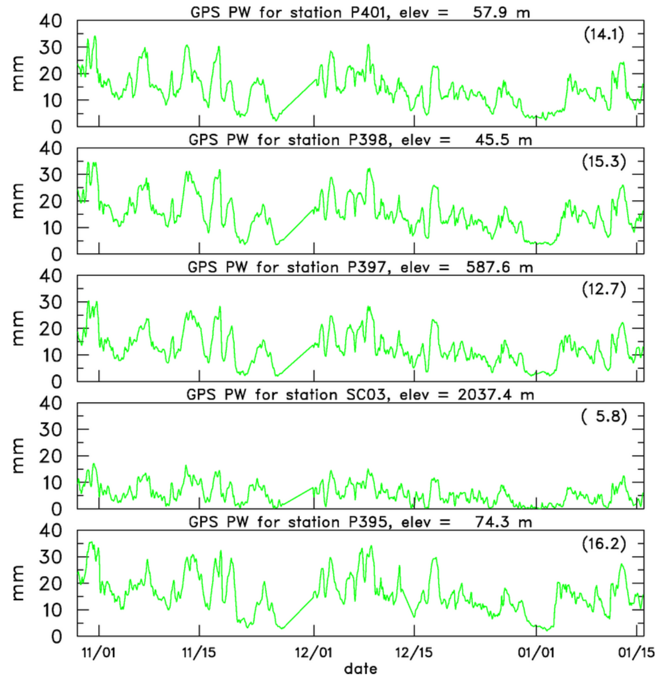

Figure 3. Time series of hourly PW (mm) data from five ground-based GPS sites within OLYMPEX domain (location of sites shown in Fig. 1). Site elevation (m) is shown in the title of each panel. Numbers in parentheses represent the period (29 October 2015 – 16 January 2016) mean. Straight line in late November 2015 indicates a period with missing data at all sites.

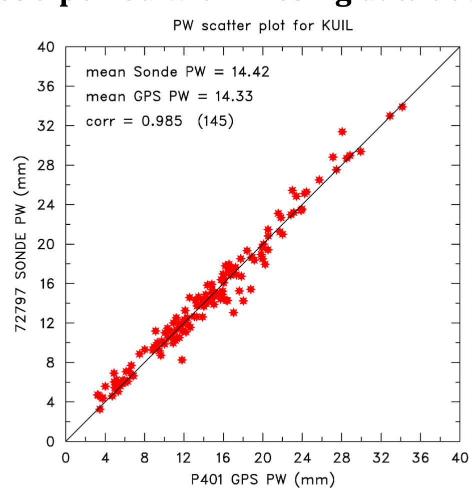

Figure 4. Scatter diagram of PW (mm) with GPS PW plotted along x-axis and sonde PW plotted along y-axis based on 145 collocated and cotemporaneous observations at KUIL. Correlation of 0.985 is significant at 99.9% confidence level.

[2] Ground-based GPS instruments sample PW in a cone with ~50 km radius above the site.

To further evaluate the PW comparison at KUIL, Fig. 5 shows a comparison over the diurnal cycle. Differences between sonde and GPS PW are less than 0.5 mm at individual hours and only 0.07 mm in the mean. These differences are well within the PW uncertainly estimates of these instruments (1-2 mm for GPS and VRS92 sondes, Yu et al. 2015). This excellent agreement gives us confidence that the humidity measurements provided by the four sites using VRS92 sondes with D3.64 software need no further humidity corrections.

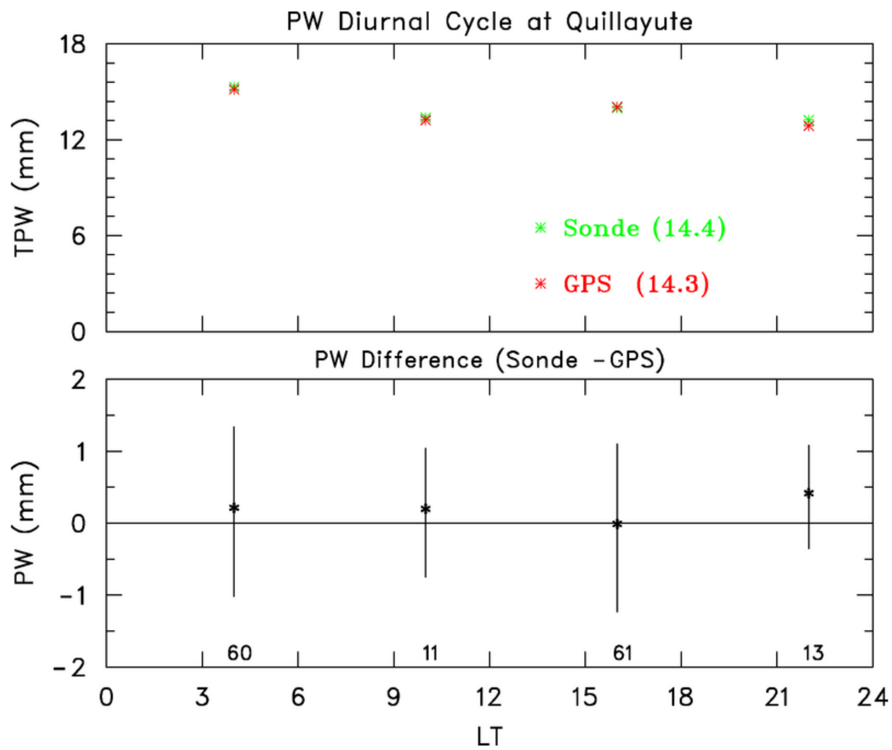

Figure 5. (top) Mean diurnal cycle of PW (mm) at KUIL with sonde (green) and GPS (red) data. (bottom) PW difference (mm) between estimates (sonde – GPS). Vertical lines indicate standard deviation of this difference. Numbers along the bottom indicate the number of comparison observations at each time. Note LT = UTC – 8.

The remaining site, KSLE, used Sippican Mark II sondes which measure humidity using a carbon hygristor. As noted in Wang and Zhang (2007), Mark II sondes exhibit a dry bias of ~5-10% in PW in low PW environments such as were present during OLYMPEX. More recent analysis (Nash et al. 2011) suggest that improvements to the Mark II hygristor may have reduced or eliminated this dry bias. Unfortunately, no GPS PW measurements were located with 50 km of the KSLE site to verify this fact and provide a basis for a correction algorithm.

Since no corrections were necessary at the four sites using VRS92 sondes or possible at KSLE, which used Mark II sondes, the L3.0 (i.e. Level 3, Version 0) dataset is merely a copy of the L2 data. Here, the L3 designation signifies that the data have been checked for sonde biases. If future analyses reveal that additional corrections of the hi-res soundings are required, these corrections will be indicated by increments in the version number.

(4) Finally, in Level 4 (L4) processing a more "user-friendly" version of the sounding dataset was created with QC flags assigned to each variable providing a measure of the data's reliability. L4 processing was performed on sites with both GTS-res and hi-res data, where L3 hi-res data were vertically interpolated to create values at uniform 5-hPa pressure intervals. Suspicious data were identified through application of both objective QC test as in Loehrer et al. (1996) and subjective adjustment of QC flags by visual inspection (Ciesielski et al. 2012) using an in-house developed visual sonde editor. While tedious, visual inspection was necessary to ensure a research-quality dataset since subtle errors in sonde data are often difficult to identify with objective procedures. An example is shown in skew-T diagram in Fig. 6 which shows a low spike in T_d near the tropopause. This artificial spike was caused by Vaisala's time-lag correction which excessively increased the humidity gradient at this level (Holger Vömel, personal communication). By flagging suspect data values, the reliable data are easily retrievable with the users deciding what level of quality is acceptable for their analyses. The definition of the QC flags used in the L4 datasets is provided in Table 2.

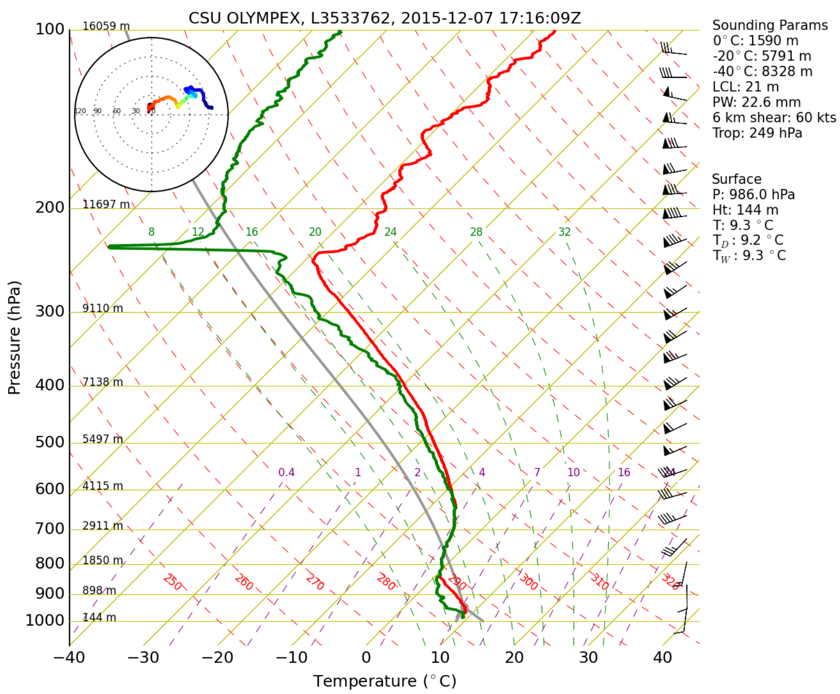

Figure 6. Skew-T plot for 07 December 2015 for NPOL site showing an artificial spike in T_d (green curve) near the tropopause level.

This second pass of QC checks and visual inspection, beyond those in L2, ensures the veracity of the data and provides yet another filter for identifying suspicious values. Note that the QC checks and visual inspection in L4 processing did not change any data values, only data quality flags. Additional details on interpolating the data to uniform pressure intervals, objective tests for assigning QC flags, and the visual editor used to expedite this processing, can be found at: www.eol.ucar.edu/projects/sondeqc/.

Flag Value	Meaning
1	parameter good
2	parameter "objectively" questionable
3	parameter "visually" questionable
4	parameter "objectively" bad
5	parameter "visually" bad
6	parameter interpolated
7	parameter estimated
8	parameter unchecked
9	parameter missing

Table 2. Convention used for the Level 4 (L4) QC flags.

4. **Data Archive**

A summary of the various sounding datasets is provided in Table 3 below.

Level 0 (L0)	Raw, original data
Level 1 (L1)	Data in common ASCII (i.e., GLS) format, native resolution
Level 2 (L2)	Data processed with ASPEN, native resolution (ASCII and netcdf formats)
Level 3 (L3.0)	Data checked for sonde biases, native resolution (ASCII and netcdf formats)
Level 4 (L4.0)	Soundings visually inspected with QC flags; hi-res data interpolated to uniform 5-hPa intervals (ASCII and netcdf formats); skew-T diagrams for all soundings

Table 3. Dataset summary and naming convection

Datasets are referenced by both a level number and version number within levels 3 and 4. For example, L3.0 signifies Level 3 version 0 of that dataset. If additional corrections are needed in the future, the version number will be incremented accordingly (e.g., L3.1). All datasets are available at the NASA Earthdata OLYMPEX archive (https://fcportal.nsstc.nasa.gov/olympex/). Also at this archive, sample fortran programs are provided to read the L1-L4 ASCII datasets.

In addition, skew-T log-P thermodynamic diagrams including various convective parameters (based on L4.0 data) have been created and are provided as part of the L4 dataset as png images contained in single compressed tar file for each site. Table 3 below shows the mean of the convective parameters for each site. Here CAPE and CIN were calculated assuming pseudoadiabatic ascent using mean thermodynamic conditions in the lowest 50 hPa. In many soundings certain parameters (CIN, LFC and EL) could not be computed due to characteristics of the sounding. In general, site-to-site parameter comparisons from this table are not very meaningful due to the different averaging periods for each site.

Site/ID	PW (mm)	CAPE (J kg^{-1})	CIN (J kg^{-1})	LCL (hPa)	LFC (hPa)	EL (hPa)
Port Hardy, BC 71109	11.6 (156)	3.7 (156)	-51.1 (21)	942.4 (156)	845.6 (21)	633.8 (21)
Env. Canada 99001	15.3 (69)	2.6 (69)	-63.8 (8)	936.2 (69)	796.3 (8)	422.3 (8)
Quillayte, WA 72797	13.6 (158)	18.8 (158)	-35.6 (23)	930.1 (158)	876.1 (24)	545.7 (24)
NPOL 99000	18.8 (96)	12.7 (96)	-37.4 (11)	939.5 (96)	831.7 (11)	502.5 (11)
KSLE 72694	14.4 (159)	3.6 (158)	-72.6 (24)	931.6 (158)	809.8 (24)	606.8 (24)

Table 3: Mean convective parameters for each site. PW signifies total-column precipitable water, CAPE – convective available potential energy, CIN – convective inhibition, LCL – lifting condensation level, LFC – layer of free convection, EL – equilibrium level. Numbers in parentheses indicate the number of soundings that went into each average.

5. Summary

During the 29 October 2015 to 16 January 2016 period, upper-air soundings were collected as part of the OLYMPEX field campaign from several dropsonde missions and five upper-air sites. In total 651 upper-air upsondes were processed and quality-controlled to produce a research-quality sounding dataset. Four of the upper-air sites used the very reliable and accurate Vaisala RS92 (VRS92) soundings. Comparison of PW between the VRS92 soundings at one site (KUIL) and collocated independent ground-based GPS observations show excellent agreement indicating the high quality of the VRS92 humidity profiles. All upsonde datasets are available at the NASA Earthdata OLYMPEX archive.

The 53 dropsondes taken as part of OLYMPEX were quality-controlled by NCAR EOL and are reported on separately. A L4 version of the dropsonde data has also been produced.

Questions regarding OLYMPEX upsonde upper-air data and its processing should be directed to Paul Ciesielski (paulc@atmos.colostate.edu).

6. References

Ciesielski, P. E., P. H. Haertel, R. H. Johnson, J. Wang, and S. Loehrer, 2011: Developing High-Quality Field Program Sounding Datasets. *Bull. Amer. Met. Soc.*, **93**, 325-336.

Loehrer, S. M., T. A. Edmands, and J. A. Moore, 1996: TOGA COARE upper-air sounding data archive: Development and quality control procedures. *Bull. Amer. Meteor. Soc.*, **77**, 2651-2671.

Nash, J., T. Oakley, H. Vömel, and W. Li, 2011: WMO intercomparisons of high quality radiosonde systems. WMO Tech. Doc. WMO/TD-1580, Instruments and Observing Methods Rep. 107, 238 pp. [Available online at: http://www.wmo.int/pages/prog/www/IMOP/publications/IOM-107_Yangjiang.pdf]

Yu, H., P. Ciesielski, J. Wang, H.-C. Kuo, H. Vömel, and R. Dirksen, 2015: Evaluation of humidity correction methods for Vaisala RS92 tropical sounding data. . *J. Atmos. Ocean. Technol.,* **32**, 397-411.

Vaisala, cited 2014: Comparison of Vaisala Radiosondes RS41 and RS92. White paper. 16 pgs. [Available online at: http://www.vaisala.com/Vaisala%20Documents/White%20Papers/Vaisala%20Radiosondes%20Comparison%20of%20RS41%20and%20RS92.pdf

Vömel, H. and Coauthors, 2007b: Radiation dry bias of the Vaisala RS92 humidity sensor. *J. Atmos. Ocean. Technol.,* **24**, 953-963.

Wang, J., and L. Zhang, , 2007: Climate applications of a global, 2-hourly atmospheric precipitable water dataset derived from IGS tropospheric products. *J. Geod.,* **83**, 209-217, doi:10.1007/s00190-008-0238-5.

Wang, J., L. Zhang, A. Dai, T. Van Hove, and J. Van Baelen, 2007: A near-global, 2-hourly data set of atmospheric precipitable water from ground-based GPS measurements. *J. Geophy. Res.,* **112**, D11107, doi:10.1029/2006JD007529.